100094567-1 Patent

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Application of:

Kay-Yut Chen

Application No: 09/944,969

Filed: August 30, 2001

For: METHOD AND APPARATUS FOR MODELING BUSINESS PROCESSES

Examiner: Morawski, Linda

Group Art Unit: 3623

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail with sufficient postage in an envelope addressed to MAIL STOP AMENDMENT, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313 on

February 01, 2006
Date of Deposit

William D. Davis
William D. Davis

MAIL STOP AMENDMENT
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

DECLARATION UNDER 37 C.F.R § 1.131

I, Kay-Yut Chen, am the named inventor of the above-referenced application. I declare that I invented the subject matter of claims 1-20 prior to the effective date of U.S. Patent Publication No. 20020169658 of Adler ("Adler") published November 14, 2002.

Adler claims priority to provisional patent application no. 60/274,328 which was filed on a purported effective date of March 8, 2001.

A paper cited as "Minimum Advertised-Price Policy Rules and Retailer Behavior: an Experiment by Hewlett-Packard", Interfaces, vol. 2, no. 35, September - October 2002, pgs 62-73 accompanies this declaration. This paper was published prior to the publication date of Adler. Page 65 references experiments performed with my invention prior to the purported effective date of Adler.

Accordingly, I declare that I invented the subject matter of the rejected claims prior to Smith's purported effective date of March 8, 2001.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Signature _____ Date 1/30/2006
Kay-Yut Chen

Minimum Advertised-Price Policy Rules and Retailer Behavior: An Experiment by Hewlett-Packard

Gary Charness • Kay-Yut Chen
Department of Economics, University of California, Santa Barbara, Santa Barbara, California 93106
Hewlett-Packard Laboratories, P.O. Box 10301, Palo Alto, California 94303-0890
charness@econ.ucsb.edu • kychen@exch.hpl.hp.com

This paper was refereed.

We tested the effects of various policy rules on retailer behavior in laboratory experiments conducted at Hewlett-Packard Laboratories. Our experimental design models the multifaceted contemporary market for consumer computer products and is quite complex, but we found that participants can make effective decisions and that their behavior is sensitive to variations in policies. Based on our results, Hewlett-Packard changed its policies; for example, it made the consequences for violations forward looking as well as backward looking. This line of research appears promising for complex industrial environments.
(*Industries: computer-electronic. Economics: experimental.*)

Manufacturers operating in the contemporary market for technology products face a daunting task in designing effective incentives for their retailers. Channels of distribution are diverse, with new channels emerging, and demand fluctuations, market exposure, advertising, stocking, and product life-cycles are uncertain. The behavior of retailers is a critical element in whether a manufacturer achieves its business goals.

Ideally, a firm would like to have a test market to determine the effects of changing its policies toward its retailers, since blindly adopting new policies in billion-dollar markets may be less than optimal. However, test markets may not be feasible in a technology market without geographic moorings; at best, test markets are expensive. While computer simulations are useful, their applicability depends on assumptions about the decisions human agents make in the field. An alternative testing method is to model retailers' choices in laboratory experiments. Economics experiments are often used in academic research to test polices; they can also be applied to business.

Such experiments have been used to examine behavior in laboratory market contexts since the publication of the seminal works of Edward Chamberlin (1948) and Vernon Smith (1962, 1964). Strong indications of external validity (applicability to the field environment) exist for behavior observed in laboratory markets, even when the environment is highly complex. For example, Vernon Smith and Charles Plott (Plott 2002, McCabe et al. 1990) pioneered smart markets to examine complex interdependent environments. In a smart market, a computerized dispatch center applies optimizing algorithms to the diverse and decentralized bids of buyers and the offers of producers and transporters to yield prices and allocations. Smith and Plott found that experimental markets can produce repeatable and predictable results. Researchers have used the methodology of experimental economics to test alternative policies in such areas as emissions trading, natural-gas pipelines, electric-power-transmission networks, transportation, and water distribution in California (Brewer and Plott 2002,

INTERFACES, © 2002 INFORMS
Vol. 32, No. 5, September–October 2002, pp. 62–73

0092-2102/02/3205/0062$05.00
1526-551X electronic ISSN

Cason 1995, Cason and Plott 1996, McCabe et al. 1990, 1991, Plott 1997, 1999, Rassenti et al. 1994).

Hewlett-Packard Company (HP) has recognized the potential of this methodology as a decision support tool; Hewlett-Packard Laboratories (HP Labs), the research arm of HP, began an experimental economics program in 1994. The firm recognizes the importance of both economic modeling and experimental methods as tools to support business decisions. Its strategy is to develop experimental models that closely mirror specific HP businesses and to then employ these models to isolate and evaluate the effects of specific policies.

Throughout its six-year history, the HP experimental economics program has performed research and developed applications in several areas, including channel management, forecasting, and electronic markets. Several HP divisions have recognized that experimental studies of the behavior of sales channels under different sets of contractual terms and business policies can provide extremely useful information before they implement such terms and policies in the field.

We developed an experimental application in the channel-management area. We are doing additional research, using game-theoretic analysis, on these issues in collaboration with John Ledyard at the California Institute of Technology. HP conducts much of its consumer business through retail channels. The distribution channels for its products include national retailers,

HP Labs began an experimental economics program in 1994.

regional retailers, mass merchant firms, clubs, and Internet retailers. Each type of retailer may have its own success metrics or business goals, which may or may not be consistent with those of HP. For example, at the time of our experimental sessions many observers felt that Internet retailers were not concerned with current profitability, as these retailers often sold to consumers at or below cost in an attempt to increase their market share. HP is concerned with the financial viability and market share of its retailers, if only because they affect its own market share and profitability.

HP uses policies to govern its relationship with its retailers, for example, return policies, price-protection policies that provide credit to the retailers corresponding to manufacturers' price fluctuations, and benefits or penalties contingent on retailer compliance with a minimum advertised-price (MAP) policy. To design effective policies consonant with its business objectives, HP must understand the implications of these policies on retailer behavior.

In this series of experiments, we studied the behavior of retailers with respect to the common industry practice of setting a minimum advertised price, a lower bound on the price a retailer can advertise for a particular product. If a retailer complies with this directive, the manufacturer typically provides it with market-development funds, which it can use to advertise the manufacturer's products. If it does not comply, it usually faces penalties. Because thousands of products are involved, MAPs are usually not enforced by legal contracts. Punishment can range from refusing to ship a product to the retailer to eliminating or reducing the amount of market development funds provided.

HP sets MAPs because it might lose market share if retailers perceived that price competition for HP products was too intense. Yet it is not clear which form of MAP (if any) is best and which enforcement policies are effective. Thus, HP wants to know what effect eliminating or modifying MAP policies would have on its market share and its retailers' profitability. An effective policy should also take into account such factors as the short life cycle of products in this market. Because it is not feasible to isolate a test market of retailers, the laboratory is an attractive alternative for investigating the impact of various policies. We conducted laboratory experiments to investigate the effects of various MAP policies on retailers' behavior and profitability, and on HP's market share.

Business Constraints and Design Limitations

We used the standard methodology of experimental economics. We brought participants into a lab and assigned them roles as retailers. They interacted with other retailers by setting prices, choosing advertising expenditures, and receiving rewards and penalties as specified by the extant policies. We gave them accurate information about the game, and told them how their

actual monetary rewards depended on their aggregate performance over the course of the session. We preserved experimental anonymity with respect to roles and payment, and we used no deception. Nevertheless, business-decision research differs from academic research. First, the experimental design went through a validation process, in which HP industry experts played the experimental game and offered feedback. Second, the business environment imposed constraints in terms of experimental design, procedures, and timetable.

HP Labs developed in-house experimental economics capabilities instead of relying on academic institutions for consultants because business considerations make such consultation impractical. Business decisions must be made in a timely fashion, even if they are made with less than perfect information. HP typically develops its potential business in three to six months, depending on the cycle of contract and policy decisions. Thus, we often design our experiments in the expectation that redesign and repetitions are unlikely, except in the most critical situations. Academic researchers generally want to establish statistical significance, necessitating replications and increasing the turnaround time.

Also, in industrial settings, it may be that no tractable theory on the research questions of interest exists, and time may prohibit developing a theoretical model that could point to specific issues to test. Because time limitations meant we could not explore the parametric space fully and because HP wished to preserve the complexity of the field environment, we tried to include as many of its features (that is, stochastic supply, demand, and delivery times, residual advertising effectiveness, and price reputation) as possible in the experiment. Our experimental environment was therefore quite complex.

We cannot vouch for the robustness of the results.

This design philosophy runs counter to standard academic experimental practice, where researchers prefer the simplest design that can encompass the modeling issues at hand. As a result, we cannot vouch for the robustness of the results. For example, if we observe some participants exploiting a policy in a certain way, we have no idea whether this behavior is an equilibrium strategy, a likely occurrence, or something that will be eliminated in the long run. However, from a business point of view, identifying such exploitation is unquestionably useful, whether or not it is the optimal strategy for a retailer. In effect, we are employing subjects to find flaws in proposed policies.

An obvious disadvantage of combining a complex design with a lack of repetition is our resulting inability to identify cause and effect. We did not control most of the many variables because of time pressure and because management does not consider it a high priority. Academic researchers may not see this approach as satisfactory; we cannot clearly attribute the findings to specific variables, as many of these were being changed simultaneously. Strictly speaking, from the standpoint of statistical analysis, we have only one observation for each session.

Nonetheless, we felt this research strategy was the most effective for obtaining the information requested in the time allocated. HP was interested in the result of changing a policy but was rather indifferent about what caused the result. The data indicate that our results are consistent with real-world observations.

Experimental Design

In our laboratory market, we attempted to model the natural setting for HP retailers. Each participant represented a retailer, while demand was computer-simulated using a model. We had heterogeneous firms interacting repeatedly in competing for consumer demand for products differentiated by price and manufacturer. Retailers made decisions about stocking, advertising, and pricing. Each (simulated) consumer considered the best price available when deciding whether to buy a product but was only aware of the products and prices to which it was exposed. A retailer's demand could also be sensitive to its reputation for pricing, relative to other retailers.

Seven differentiated retailers interacted in each of our sessions. They were intended to represent national firms, PC Direct/Mail Order companies, mass merchants, clubs, and Internet retailers. PC Direct companies are ones that sell HP printers with their PCs.

Each retailer chose a price for each product in each period and competed for some percentage of the potential market for the products. Most firms could increase this percentage by advertising, although each type of retailer had a maximum exposure percentage and advertising yielded diminishing marginal returns. Most retailers also had to make inventory decisions, with the cost of holding excess inventory balanced against a negative reputation if a retailer failed to meet most of the demand for a product. The timing of deliveries to the retailers was stochastic.

We computer-simulated consumer demand using a random utility multilevel logit model (Dubin 1998, McFadden 1976) adapted to the HP environment by Steven Gjerstad and Jason Shachat. This model treats each product as a collection of attributes (such as price, brand, retailer, speed, and memory). When assessing a potential product choice, each consumer assigns a different weight to the value of each attribute, and the model adds these values together to determine that consumer's score for the product. The probability that the consumer purchases a product increases with this score, and the probability that any one product is selected is the estimated market share of that product.

HP was interested in the result of changing a policy but indifferent about what caused the result.

The stochastic market size lies within a range known to the retailers, who also receive a signal that further limits this range at the beginning of a period.

Retailers can sell products offered by HP and by competing manufacturers. These products vary by retailer cost and by manufacturer policies on product returns and advertising. We evaluate different retailers using diverse measures that reflect the contemporary business goals of the different categories of retailer. These measures include various combinations of gross profit, net income, revenue, and GMROII (GMROII is based on the product of revenue and the ratio of gross profit to total inventory value for the past four periods; this is a common performance measure in this industry). The model incorporates product obsolescence through a life-cycle assumption—some products get phased out and others take their place, with retailers receiving notice five periods in advance.

Inventory control is a crucial aspect of the natural retailer environment. Most retailers (although not all) need to stock products to be able to sell them. However, it is usually costly to carry excess inventory. In addition, while a retailer may place an order for products, the actual shipment date is uncertain. Further, supplies may be short at any particular time. Retailers must consider all of these factors when making stocking decisions; a retailer who cannot meet existing demand develops a negative reputation for service, which negatively affects subsequent demand.

Finally, advertising clearly affects demand and must be considered, particularly because advertising policy is the control variable in the experiment. A retailer has some minimum level of market exposure even without any advertising. However, advertising increases market exposure in a nonlinear fashion, until it saturates the market for the retailer. While a firm may be free to advertise any price it likes, violating manufacturer mandates concerning minimum advertised price jeopardizes the advertising funds potentially available from the manufacturer. Manufacturers employ several schemes to punish violations.

The natural market is very complex and even chaotic, with new types of retailers growing in importance. Planners within manufacturing firms must somehow formulate policies that take important marketplace features into account without making decisions so difficult that the results are arbitrary.

Experimental Procedure

We conducted our first set of sessions in September 1999. We used the insights obtained in September to modify our design for our second set of sessions in February 2000. (Detailed experimental instructions are available upon request; we omit the fine detail of our calibrations and models to protect intellectual property.)

We recruited participants by sending an e-mail message to Stanford University interest groups. Most of our subjects turned out to be graduate students. Because of the complexity of the environment and the need for participants to make several decisions each

Retailer #	Must stock?	Can advertise?	Minimum % market exposure	Maximum % market exposure	Evaluation method
1	yes	yes	30	100	70% GMROII, 30% Net income
2	yes	yes	30	100	Gross profit
3	yes	yes	30	70	70% GMROII, 30% Net income
4	yes	yes	30	70	70% GMROII, 30% Net income
5	yes	yes	30	50	100% GMROII
6	yes	no	40	40	70% GMROII, 30% Net income
7	no	yes	10	30	Revenue

Table 1: In our September experiments, participants played the roles of seven very different types of retailers. They differed in many aspects from their reach in the market (min/max percent market exposure) to whether they stocked and held their own inventories (some retailers fulfilled orders through a third party and held no inventory).

period, our initial sessions were quite lengthy (we have now developed a design that facilitates much shorter sessions).

In establishing pay rates for participants, we tried to calibrate expected earnings to about $18 per hour (including a show-up fee), and actual earnings ranged from $10 to $25 per hour. However, we could not make any guarantees about pay, and the time requirement made it rather difficult to fill the sessions. Participants were paid a show-up fee of $25, and their remaining earnings were based on their profitability. We used a dollar conversion rate that varied by the type of retailer.

The participant-retailers viewed information on a series of six screens:

(1) The order screen offered them an opportunity to make purchases and listed past-period pricing and margins for each retailer, how much was spent on advertising for each product in the upcoming period, and inventory and ordering information.

(2) The advertising screen again presented pricing and inventory information and also stated the amount available for advertising. Participants chose advertising expenditure for each product. Advertisements ran four periods after the space was reserved, although the retailer chose the advertised price in the period that the advertisement appeared. There was a two-period lag between the choice of advertising expenditures and the appearance of the advertisement (except for retailer 7, who had no lag). Retailers could advertise only with the advertising funds provided in an initial endowment and later supplemented by manufacturer funds based on product purchases.

(3) Retailers chose selling prices (these were also the advertised prices for that period) on the pricing screen, which again listed pricing and inventory information and also indicated any pricing restrictions. There were no restrictions on the price per se; if no advertisement appeared for a product in a period, no MAP violation would occur, regardless of the selling price.

(4) The price-control-and-ad screen showed the advertising funds earned from the shipments received in that period, the amount lost in that period because of a MAP violation, and the number of periods remaining in the MAP penalty.

(5) The supply, demand, and return screen showed the retailer's demand for each product for the period. If units had been ordered but supply was rationed, this was indicated. If a stock-out penalty (for servicing less than 50 percent of the experienced demand) was in force, this was indicated, along with the number of periods remaining for this penalty. Retailers could return products, up to a limit of 6 percent of cumulative shipments received. Because of the advertising lag, we did not begin demand until period 5.

(6) The earnings summary screen evaluated the retailer's performance for that period and for the entire session, using the appropriate metric.

Our markets had seven retailers of various types. Sessions lasted about three hours. Each person was seated at a computer in a carrel separated from others by dividers so that participants could not observe others' decisions.

The retailers were very different (Table 1). For example, a club retailer (Number 6) doesn't advertise,

while an Internet retailer (Number 7) has a small potential market share, keeps no stock, and has only one performance metric-revenue. We used various rates for converting participants' experimental earnings into the actual dollars we paid them, reflecting the heterogeneity of types of retailers. We also differentiated products with respect to cost and levels of demand. This experimental setup was used to test several penalty strategies (Table 2).

In the February experiments, we lengthened the sessions to seven hours, including a two-hour training period. We made a number of design changes; for example, we restricted the number of products that retailers could advertise in a period, we included a factor for historical price reputation, and we slightly modified the performance measures for the retailers. We had found that the penalties for MAP violations were ineffective near the end of the experiments (or life cycles) in the September sessions, and so we made them partially retroactive in February (Table 3).

In the training segment of each session, we presented an overview of the experiment. We summarized the mechanics involved in making choices and the effects of these choices on retailer performance. We also discussed stocking issues, service levels, pricing, advertising and demand, advertising funds, and product life cycles. With respect to product life cycles, we told participants that we would replace two products during the session and that we would notify them of this change five periods in advance. We also described the evaluation methods in some detail. We gave each participant a chart that illustrated the sensitivity of his or her own demand to advertising expenditures.

We also covered the MAP violation penalties, providing retailers with a chart of the penalties for each product in that session. Possible penalties in the September sessions included pulling products (preventing a retailer from receiving further shipments), suspending advertising funds for a number of periods, and withdrawing advertising funds for the current period. In one session, we linked all HP products, so that a violation on one product triggered penalties on all. In the February sessions, we based some penalties on net shipment value and revenue.

After a question-and-answer period in each session, participants played some practice rounds to further familiarize themselves with the mechanics involved in the experiment. We answered individual questions during this practice phase as well, and then we proceeded with the experiment (Table 4).

The simulation determined demand after the first three decisions, and retailers chose their returns after observing this demand. Each retailer made these four decisions for each of eight products in each of seven to 11 periods: how many units to order, how much to spend on advertising, what prices to charge, and how many units to return to the manufacturer.

Results

In the September sessions, we used penalties that primarily applied to future periods. We varied these penalties for products 1 through 4, the control products. We kept penalties for the remaining products constant across treatments; for products 5 and 8, a violation meant losing four periods of ad funds, while for products 6 and 7, a violation meant being fined the current period's ad expense.

Because we observed that forward-looking penalties became less effective as we neared the end of each September session, for the February sessions we made the penalties also retroactive for some number of periods. Again, we varied the penalties for products 1 through 4 and held the penalties constant for products 5 through 8. In one session, we imposed multiperiod penalties on MAP violations for products 1 through 4. In two other sessions, we removed price restrictions for either products 1 through 4 or for only products 1 and 4 (product 1 [or product 4, its life-cycle replacement] has the largest market share). We ran 20 periods in one session, eight in a second session, and 12 in a third session.

The September sessions (Table 5) differed with respect to the penalty for a MAP violation for products 1 through 4 with the violation penalty for products 5 through 8 kept constant across sessions. In the February sessions (Table 6), we imposed the restriction that a retailer could advertise at most two products in any one period.

Before moving to our analysis, we caution against imputing statistical significance to our results because

Product	September (1) Penalty	September (2) and (3) Penalty	September (4) Penalty
1	Pulled	4 periods ad funds	12 periods ad funds
2	Pulled	4 periods ad funds	12 periods ad funds
3	Pulled	4 periods ad funds	12 periods ad funds
4	Pulled	4 periods ad funds	12 periods ad funds
5	4 periods ad funds	4 periods ad funds	4 periods ad funds
6	Current period ad expense	Current period ad expense	Current period ad expense
7	Current period ad expense	Current period ad expense	Current period ad expense
8	4 periods ad funds	4 periods ad funds	4 periods ad funds

Table 2: There were four types of treatments in our September experiments. They differed in the MAP violation penalties for products 1 through 4, which represent Hewlett-Packard products. (2) and (3) also differed in the number of products to which the penalties applied. In (3), if the advertised-price restriction is violated for any of products 1 through 4, the MAP violation penalty applies to all of these products.

of the interdependence of the observations in each session. Individual sessions varied considerably, further weakening statistical comparisons. Nevertheless, we see some patterns in the data.

The overall market share of the control products was only slightly reduced by having less severe MAP violation penalties for these products. In both sets of sessions, a comparison of the harsher penalties with aggregated gentler penalties shows that the control-product market share is about 10 percent higher with the more severe penalties. While this small difference may seem surprising, it may be the result of a correlation between the pricing of the control products and the other products in any one session. Thus, HP does better with harsher penalties but only slightly.

In both September and February, we found that retailer margins were higher with the more severe penalty. This was true for both sets of products even though we held the penalties for the noncontrol products constant across treatments. This finding suggests that the retailers' pricing decisions for all goods are sensitive to the nature of the penalties for violating MAP on the control products.

In September, the average margins were about 20 percent higher when a violation led to products being permanently pulled from the retailer (for reference, we set the price restrictions so that the average margin at

Product	February (1) Penalty	February (2) Penalty	February (3) Penalty
1	3% + 3%*	No penalty	No penalty
2	3% + 3%	No penalty	3% + 3%
3	3% + 3%	No penalty	3% + 3%
4	3% + 3%	No penalty	No penalty
5	4 periods ad funds, starting this period	4 periods ad funds, starting this period	4 periods ad funds, starting this period
6	Current period ad expense	Current period ad expense	Current period ad expense
7	Current period ad expense	Current period ad expense	Current period ad expense
8	4 periods ad funds, starting this period	4 periods ad funds, starting this period	4 periods ad funds, starting this period

Table 3: There were three types of treatments in our February experiments. They differed in the MAP violation penalties for products 1 through 4, which represent Hewlett-Packard products. 3% + 3% means lose three percent of net shipment value for the past four periods plus three percent of revenue for the current period and the next three periods.

Choose	Observe	The action affects
Number of units to order	Approximate number of customers	Stock available later
	Current inventory position	Service levels
	Buying prices (past service-level)	
	Past selling prices	
Advertising	Advertising budget offered	Later demand
	Approximate number of customers	
	Current inventory position	
	Buying prices (past service-level)	
	Past selling prices	
Selling prices	Approximate number of customers	Prices charged customers
	Current inventory position	Current period demand
	Competitors' last period prices	
Number of units to return	Demand	Stocking levels
	Stock remaining	Service levels

Table 4: The participants made several decisions and observations in the course of the experiment. These variables were summarized and printed on reference sheets, which were handed out to participants during the experiments.

	Control products (1–4)		Other products (5–8)	
MAP violation penalty (products 1–4)	Average margin	HP share of market	Average margin	HP share of market
Lose 4 periods ad funds	0.10 (0.02)	55%	0.16 (0.03)	45%
Lose 4 periods ad funds (linked)	0.11 (0.04)	55%	0.16 (0.03)	45%
Lose 12 periods ad funds	0.08 (0.04)	41%	0.11 (0.03)	59%
Aggregated ad funds penalties	0.10 (0.03)	50%	0.14 (0.04)	50%
Pull the product	0.12 (0.01)			54%

Table 5: In the September experiments, the pull-the-product penalty was the most effective with the highest margin observed compared to the ad-funds penalties.

	Control products (1–4)		Other products (5–8)	
MAP violation penalty (products 1–4)	Average margin	HP share of market	Average margin	HP share of market
No MAP, products 1–4	0.03 (0.01)	56%	0.06 (0.03)	44%
No MAP, products 1&4	0.00 (0.09)	49%	0.04 (0.10)	51%
Aggregated no MAP	0.02 (0.06)	53%	0.05 (0.07)	47%
Backward/forward penalty (3% + 3%)	0.11 (0.04)	59%	0.14 (0.05)	41%

Table 6: In the February experiments, we found the new 3% + 3% penalty to be as effective as the previously tested (September experiments) penalties. It also maintained substantially higher margins and market share compared to scenarios to which it was not applied.

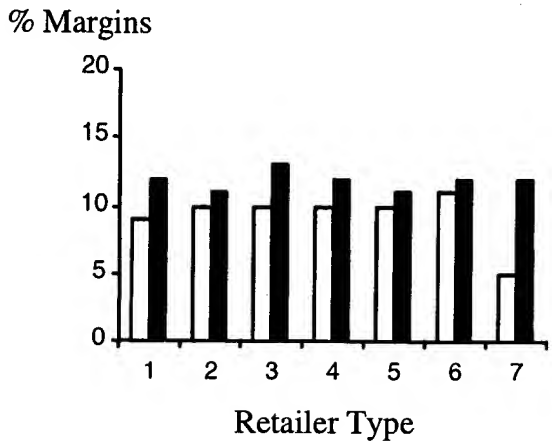

Figure 1: The September experiments showed that retailer margins for the control products are higher with the pull-product penalty (black) than with the ad-funds penalties (white).

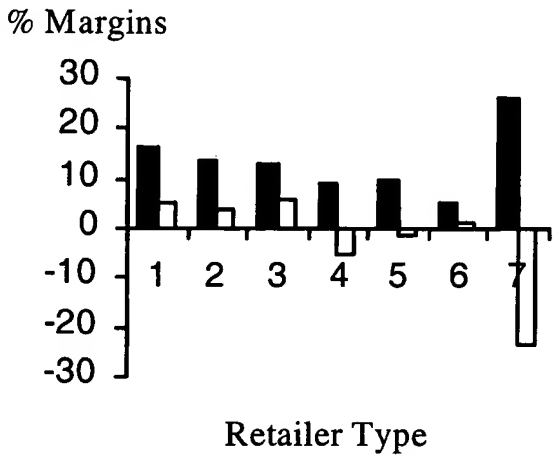

Figure 3: The February experiments showed that retailer margins are substantially higher for the control products with MAP penalties (black) than without MAP penalties (white).

the restricted price was 10 to 13 percent for the control products (Figure 1) and 17 to 20 percent for the other products (Figure 2)). If we were to assume the independence of each observation, this difference would be statistically significant at $p = 0.04$ (one-tailed test).

The margins were always lower for retailers when the penalty for violating a control-product MAP was only temporary. Even though the penalties vary only for the control products, this is true for all 14 comparisons.

The difference in margins was even more pronounced in the February sessions (Figures 3 and 4). The margin with MAP is significantly higher than the margin for the combined sessions without MAP at $p = 0.002$ (one-tailed test).

It is apparent that the margin for individual retailers on all products is robustly higher with strict penalties for MAP violations.

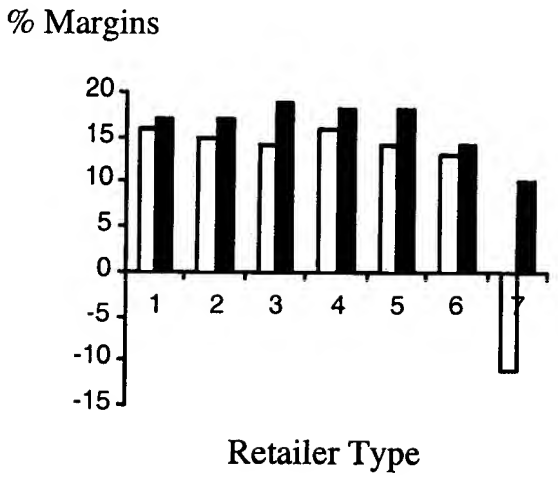

Figure 2: The September experiments showed that retailer margins for other products are higher with the pull-product penalty (black) than with the ad-funds penalties (white).

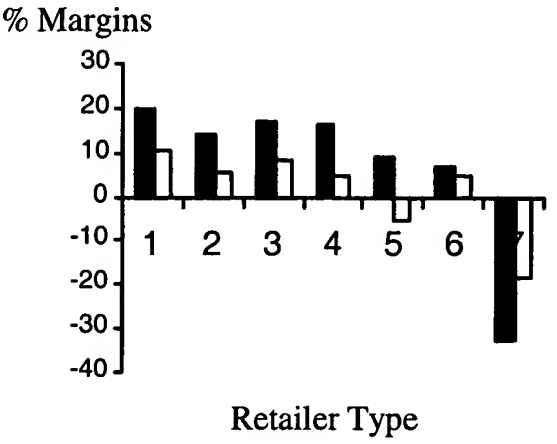

Figure 4: The February experiments showed that retailer margins are substantially higher for the other products with MAP penalties (black) than without MAP penalties (white).

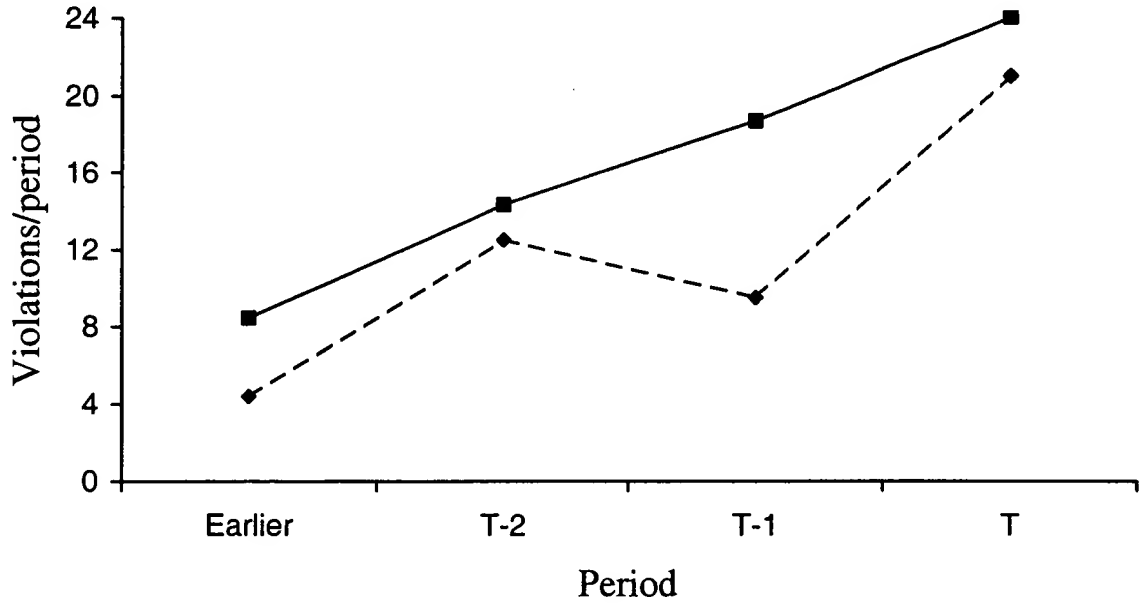

Figure 5: MAP violations per period in the September experiments showed an upward trend under both the pull-products penalty (dashed line) and the ad-funds penalties (solid line). T represents the final period of a session, T-1 the penultimate period, and so forth.

In the September sessions, we used an exclusively forward-looking violation. We observed a pattern in the violation rate over time: close to the end of the experiment, every retailer violates MAP substantially more. A forward-looking penalty should (and did) have diminishing effectiveness as a product approaches the end of its life cycle (Figure 5).

We see a positive time trend in the number of violations per period, as there are more violations as the end of the life cycle approaches. In the February sessions, we introduced a violation penalty with a retroactive component.

We found that the frequency of violations was related to the form of MAP imposed. We also found that retailers (particularly mid-sized retailers) did not fare as well without MAP, as their margins were distinctly smaller; interestingly, removing the MAP on some products affects the margins for both those products and for the others. This calibration suggests that equilibrium prices may well be below the price floor. Based on our results, HP felt it would be best to continue some form of MAP.

We were also able to detect weaknesses in the design and enforcement of several advertised-price policies; this led HP to revise the policies it implemented. For example, retailers may carry several different HP products. One proposed enforcement policy would link these products, so that a violation on any individual product would trigger penalties on all of them. When we tested this policy, we found that retailers who decided to violate the MAP on one product would often violate the MAP on all the linked products. As a result, HP decided not to implement a linked-product MAP design.

In addition, in our first set of sessions, we identified a problem with respect to MAP and product life cycle. Initially, we tied MAP penalties to future shipments and future market-development funds for the product at issue. However, we found that the violation rate increased toward the end of the life of a product. Retailers correctly perceived that forward-looking penalties would have little effect late in a product's life. Because of this, HP decided to adopt a completely different enforcement policy, which we validated in our second set of sessions (Figure 6). This new policy is retroactive as well as forward looking, so that retailers cannot

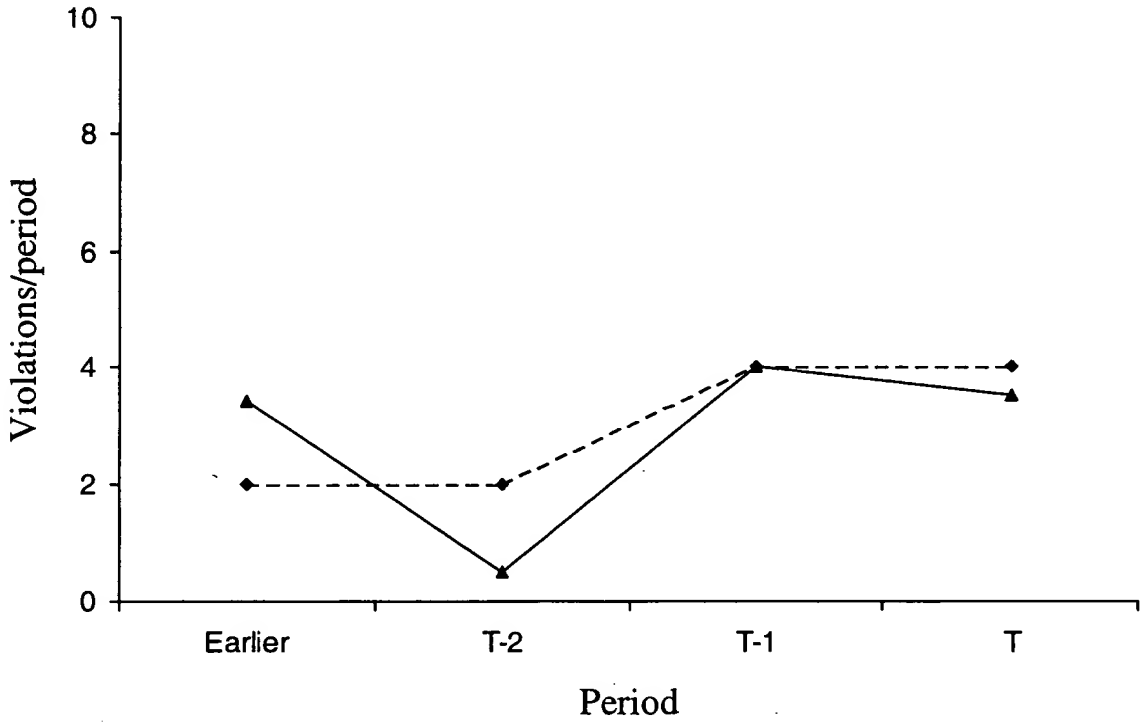

Figure 6: The MAP violations per period in the February experiments no longer showed an upward trend under either the pull-products penalty (dashed line) or the ad-funds penalties (solid line). Here we see no real time trend. This approach seems to have been effective in reducing the violation rate near the end of a session or life cycle.

escape penalties even if they violate MAP at the end of a product's life.

Discussion

Our aim in this research was to examine the effect of various penalties for violating MAPs on retailer behavior and on HP's market share. Retailer margins appear to be inversely related to the severity of the penalties for violating MAPs. Changing the penalties for the control products seemed to have only modest effects on their market share. We learned that a penalty that links products has a serious flaw, and HP decided not to use such penalties. We also found that purely forward-looking penalties led to a pattern of increasing violations as products approached the ends of their life cycles and that including a retroactive component in the penalties seemed to be effective in reducing or eliminating this effect. HP has subsequently developed a new design based on these results, introducing backward-looking penalties to counter the life-cycle effect and eliminating linked-product penalties.

Our study has many limitations, and our methodology is still evolving. We learned some lessons that might be useful to others who wish to apply experimental methods in industrial applications. A firm may wish to match its business environment as closely as possible in an experiment, but doing so may require a design that is too complex for conventional analysis. In practice, the researcher and the industrial client may need to negotiate the details of the experiment.

In addition, complex experiments may take a long time to run; we found that recruiting participants for our longer sessions was difficult. We recommend that prospective experimenters keep recruiting issues in mind. Streamlining the decision path should be helpful. We are working on new interface functionality, which should reduce the time needed for a session and

may ameliorate the recruiting problem. Our design can accommodate a variety of retailer types, an important factor given the changing markets for technology products.

The experimental approach seems promising for business enterprises wishing to evaluate the effects of policy changes, even in complex market environments. Sometimes even a limited examination of potential strategies is useful and can produce surprising dividends. Our associates in the Hewlett-Packard product divisions recognized the value of our experimental results for making business decisions and setting policies.

Acknowledgments

We acknowledge the valuable support provided by Kemal Guler, Shailendra Jain, Fereydoon Safai, and Jerry Shan (in particular, Jerry's work on the demand model) at HP Labs, and Craig Artherholt, Jacky Churchill, Richard Deep, Alex Espalin, Alan Maybruck, Sheila McKay, and Robin McShane at HP business divisions. We also gratefully acknowledge the helpful and extensive comments of two anonymous referees.

References

Brewer, P., C. Plott. 2002. A decentralized, smart market solution to a class of back-haul transportation problems: Concept and experimental test beds. *Interfaces* **32**(5) 13–36.

Cason, T. 1995. An experimental investigation of the seller incentives in EPA's emission trading auction. *Amer. Econom. Rev.* **85**(4) 905–922.

——, C. Plott. 1996. EPA's new emissions trading mechanism: A laboratory evaluation. *J. Environ. Econom. Management* **30**(2) 133–160.

Chamberlin, E. 1948. An experimental imperfect market. *J. Political Econom.* **56**(2) 95–108.

Dubin, J. 1998. *Studies in Consumer Demand—Econometric Methods Applied to Market Data.* Kluwer Academic, Boston, MA.

McCabe, K., S. Rassenti, V. Smith. 1990. Designing 'smart' computer-assisted markets. V. Smith, ed. *Papers in Experimental Economics.* Cambridge University Press, New York, 678–702.

——, ——, ——. 1991. Experimental research on deregulated markets for natural gas pipeline and electric power transmission networks. *Res. Law Econom.* **13** 161–189.

McFadden, D. 1976. Quantal choice analysis: A survey. *Ann. Econom. Social Measurement* **5** 363–390.

Plott, C. 1997. Laboratory experimental testbeds: Application to the PCS auction. *J. Econom. Management Strategy* **6**(3) 605–638.

——. 1999. Policy and the use of experimental methodology in economics. L. Luini, ed. *Uncertain Decisions Bridging Theory and Experiments.* Kluwer Academic Publishers, Boston, MA, 293–315.

Rassenti, S., V. Smith, K. McCabe. 1994. Designing a real time computer-assisted auction for natural gas resources. W. Cooper, A. Whinston, eds. *New Directions in Computational Economics.* Kluwer Academic Publishers, Boston, MA, 41–54.

Smith, V. 1962. An experimental study of competitive market behavior. *J. Political Econom.* **70**(2) 111–137.

——. 1964. The effect of market organization on competitive equilibrium. *Quart. J. Econom.* **78**(2) 181–201.

Fereydoon Safai, Project Manager, Decision Technology Department, Software Technology Laboratory, Hewlett-Packard Labs, PO Box 10301, Palo Alto, CA 94303-0890, writes: "The work described in the paper has been transferred to and adopted by the Consumer Product Organization of Hewlett-Packard. Major policy decisions were made based on the information provided by Dr. Chen's and Dr. Charness's research. Jacky Churchill, Vice President and General Manager at the time of the research, wrote in an internal memo, 'The beauty of the model is its ability to allow us to "test" a number of different variations, and see the effects, without creating a "disturbance" in the marketplace.' In the past year or so, this research has also expanded to cover other policy areas and has provided substantial value to the Consumer Product Organization on many occasions."